

Selena (Ruiqi) Ge

Data Scientist | Researcher | Future Mathematician | Cat Person

9245 Regents Road Unit M223, La Jolla, CA, 92037 | (+1)858-967-3988 | (+86) 18802418105 | sege@ucsd.edu

Welcome to My Homepage: <https://selenageruiqi.github.io/> | Github: <https://github.com/SelenaGeRuiqi>

EDUCATION

University of California San Diego, Master of Science

Sep 2024 – May 2026(Expected)

Major: *Data Science*.

Cumulative GPA: 4.0/4.0

New York University Shanghai(NYUSH), Bachelor of Science

Sep 2020 – May 2024

Major: *Mathematics*

Minor: *Business and Finance*

Cumulative GPA: 3.84/4.0

Major GPA: 3.95/4.0

Honors: Major Honors in Arts and Sciences, NYU Shanghai Excellence Award, NYU Founder's Day Award, Dean's list for 4 academic years, Latin Honors: Cum laude.

Language & Tools: Python(Pytorch, Tensorflow, Pandas, NumPy,Scikit-learn), R, MATLAB, LaTeX, SQL, Tableau, etc.

Research Interest: Reinforcement Learning(RL), Diffusion Models, Stochastic Modeling, Causality, Machine Learning(ML).

Publications and Symposiums:

1. Fan Feng, **Selena Ge**, Minghao Fu, et. Ada-Diffuser: Latent-Aware Adaptive Diffusion for Decision-Making. [OpenReview](#) Preprint. Under review for NeurIPS 2025.
2. **Ruiqi Ge**, Jiaqi Liu. Spillover Effects of Virtual Tours on Tourism in China during the Pandemic Era.
 - The 18th International Research Symposium on Service Excellence in Management ([QUIS18](#)). [Conference Publication & Presentation]
 - NYUSH 2022 Dean's Undergraduate Research Fund ([DURF](#)) Symposium. [Presentation]

WORK EXPERIENCE

HP, Corporate Venture Capital & Business Analytics program | Business Analyst | New York, US

Jul 2022 – Sep 2022

- Independently developed a quantitative evaluation system to identify high-potential startups offering hybrid work platforms during pandemic; conducted data scraping from 200+ startups using Selenium and data cleaning with Pandas.
- Performed statistical modeling and feature selection using SPSS and R (Pearson correlation, factor analysis, multicollinearity diagnostics); Constructed a multi-factor scoring model, assigning weights based on variance contributions and benchmarking against HP's historical investments.

Neusoft, Medical IT Business Division | Software Engineer Intern | Shenyang, CN

Jan 2022 – Feb 2022

- Co-led the on-site IT project for large-scale operational data management in top-30 nationally ranked digitized hospital.
- Developed a business intelligence model for multidimensional dynamic visualization of hospital operation metrics; tackled key challenges in metric development and contributed substantially to the enhancement of data visualization.
- Engaged in optimizing data collection and warehousing modeling methods focusing on the technical aspects of OLTP databases unique to hospital business systems, which significantly improved the efficiency of data operation by 25%.

NYUSH, Academic Resource Center | Lead Learning Assistant | Shanghai, CN

Sep 2021 – May 2022

- Hosted biweekly review sessions for over 60 students on *Linear Algebra* to enhance their comprehension of new concepts and provided after-class tutoring for 4 hours every week.
- Communicated regularly with the Academic Department managers; co-facilitated academic workshops, training, and discipline-related projects for all colleagues weekly; received Excellent Communication Award.

SELECTED RESEARCH

[NeurIPS 2025 Under Review] Ada-Diffuser: A Causal Diffusion Framework for Latent Identification in RL | Co-First Author

Direction: Causal Representation Learning, RL, Diffusion Models; Mentor: Prof. Biwei Huang Mar 2025 – Present

- Contributed to the theoretical foundation of Ada-Diffuser, a novel framework that integrates latent variable identification with causal diffusion modeling for reinforcement learning.
- Derived identifiability conditions under minimal temporal observations and supported the formulation of latent-contextual POMDPs using structural causal models (SCMs).
- Provided key insights on integrating diffusion models into sequential decision-making by modeling autoregressive causal processes with time-dependent latent factors.

Enhancing Cross-Cultural Toxic Speech Detection with Concept Bottleneck LLMs | *Independent*

Direction: NLP, Interpretability, Fairness; Mentor: Prof. Lily Weng Oct 2024 – Present

- Initiated the idea of extending Concept Bottleneck LLMs (CB-LLM) to multilingual toxic speech detection, targeting fairness and interpretability across cultural and linguistic contexts.
- Formulated the theoretical foundation for a four-stage pipeline (concept generation, automatic scoring, bottleneck alignment, and prediction), leveraging SimCSE and ACS for cultural concept grounding.
- Contributed to the design of experiments and evaluation strategy, helping guide implementation efforts and ensure alignment with ethical and explainable AI principles in real-world moderation tasks.

[Undergraduate Senior Thesis] Stochastic Differential Equations(SDE) for Large-scale Machine Learning: Applications on Least Squares Stochastic Gradient Descent and Score-Based Generative Models | *Independent*

Direction: Stochastic Process & Deep Learning; Mentor: Prof. Mathieu Lauriere & Prof. Roberto Fernandez Feb 2024 – May 2024

- Clarified the connection between Stochastic Differential Equation (SDE) and two cutting-edged diffusion models, demonstrating how score-based methods can reverse the noise addition process inherent in diffusion models.
- Carried out both detailed mathematical deduction and numerical experiment to confirm the efficacy of Denoising Diffusion Probability Model(DDPM) and Score Matching Langevin Dynamics (SMLD) within the SDE framework.

The Evolution of Shanghai 50 ETF Option in the Covid Era Based on Stochastic Models and ML | *First Author*

Direction: Stochastic Modeling & ML; Mentor: Prof. Roberto Fernandez & Prof. Loucas Pillaud-Vivien Mar 2023 – Jan 2024

- Simulated the evolution of China's stock market performance numerically from 2013 to 2023 based on the enhanced Black-Scholes Model and achieved a remarkable alignment with actual market data using Python and VBA.
- Constructed a Stochastic Gradient Descent (SGD) model based on stochastic modified equations (SME), which improved the simulation performance and efficiency by 65% and demonstrated effectiveness for both small-scale and large-scale datasets.
- Defined a Stochastic Differential Equation (SDE) model through Python which is consistent with the small step-size limit of the SGD recursion, and greatly enhanced the simulation process by using the Euler-Maruyama method.

Bayesian Modeling of Information Spillovers in Post-Pandemic Travel Decisions | *Co-first Author*

Direction: Bayesian Modeling, Behavior Economics; Mentor: Prof. Jiding Zhang Jan 2022 – Jun 2023

- Redefined an innovative Bayesian model to simulate tourists' decision-making results under 8 different circumstances.
- Constructed an instrumental variable regression model to quantify the relationship between 5 key factors using R; affirmed the model's robustness at 95% confidence level via instrument relevance and exogeneity tests using Stata.
- Presented research findings both at the 2022 Fall NYUSH DURF Symposium and at the *18th International Research Symposium on Service Excellence in Management* (QUIS18 Conference) at Vietnam University in the summer of 2023.

Decision-making Project: Should Shanghai Disney Implement the FastPass System? | *Leader*

Direction: Business Analytics; Mentor: Prof. Grace Haaf Nov 2022 – Dec 2022

- Predicted the net present value before and after implementing the FastPass system using simulation and regression, obtaining R-squared over 90%; tested relationship between FastPass implementation and influential factors via A/B testing.
- Estimated the times of visits from the sentiment of online reviews via Natural Language Processing; explored the customer segmentation by K-Means and Decision Tree, and visualized the result using Tableau.